GUJARAT TECHNOLOGICAL UNIVERSITY
BE-MINOR - SEMESTER–V EXAMINATION – WINTER 2022

Subject Code:115AG01   **Date:22-02-2023**
Subject Name:Introduction to Data Analytics
Time:10:30 AM TO 01:00 PM   **Total Marks:70**
Instructions:
1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Simple and non-programmable scientific calculators are allowed.**

			Marks
Q.1	**(a)**	Define the term data science. Enlist the various sectors using data science.	**03**
	(b)	List out various applications of data science. What is the role of the next generation data scientist?	**04**
	(c)	Write a python program to plot bar chart with suitable dataset. Explain various attributes of the bar chart.	**07**
Q.2	**(a)**	Define the term: Regression analysis.	**03**
	(b)	Illustrate the technique of missing value treatment with an example.	**04**
	(c)	What is Principle component analysis? Explain how to perform it in detail.	**07**
		OR	
	(c)	What do you mean by categorical variables? Explain the various techniques of dealing with categorical variables.	**07**
Q.3	**(a)**	Define the terms: Data, Data frame, Dataset	**03**
	(b)	Differentiate between feature selection and feature extraction.	**04**
	(c)	Explain K-means clustering algorithm. What is the major drawback of it?	**07**
		OR	
Q.3	**(a)**	List out various libraries of python used for data analytics.	**03**
	(b)	Explain Logistic regression with illustration.	**04**
	(c)	Define Bayes theorem. Explain Naïve Bayes classifier in detail.	**07**
Q.4	**(a)**	Enlist the types of graphs. State how to choose it?	**03**
	(b)	Enlist the various feature selection methods and explain any one in brief.	**04**
	(c)	How to perform data preprocessing? Give suitable example to detect and replace null values in python.	**07**
		OR	
Q.4	**(a)**	Explain visualization on complex dataset.	**03**
	(b)	What is dimensionality reduction? Justify the importance of it.	**04**
	(c)	What is EDA? Explain the concept by giving suitable example.	**07**
Q.5	**(a)**	What do you mean by multiclass classification?	**03**
	(b)	Outline the main advantage of visual data exploration.	**04**

(c) Consider the following case study. Determine the suitable model used for performing data analytics on the same and provide justification for the selected model. **07**

Case Study: Stock market prediction is the act of trying to determine the future value of company stock or other financial instrument traded on an exchange. The successful prediction of a stock's future price could yield significant profit.

OR

Q.5	**(a)** What do you mean by data visualization?	**03**
	(b) What do you understand by predictive analytics?	**04**
	(c) Explain any one application of data analytics with its detailed process and suitable dataset.	**07**